DCC phenotype tagging details

Updated 05/10/2021

This page provides additional information about finding the results of the [DCC's phenotype tagging project](#) [1]. All tags have been mapped to a [UMLS](#) [2] Concept Unique Identifier (CUI), which is required for identifying the tagged variables on dbGaP.

Contents

- [Instructions for identifying tagged variables on dbGaP](#) [3]
  - [Entrez search](#) [4]
  - [Faceted search](#) [5]
  - [Variable page](#) [6]

- [The 65 phenotype tags mapped to UMLS concepts](#) [7]

[Back to top](#) [8]

Instructions for identifying tagged variables on dbGaP

Please see our [YouTube demo](#) [9] showing how to search for tagged variables on dbGaP. The following are examples of different methods to search for tagged variables.

[Back to top](#) [8]

Entrez search

- In your web browser, visit the [dbGaP Entrez advanced search page](#) [10]
  - You can navigate here by clicking on the Advanced link under the Entrez search box at the top of the dbGaP home page
- In the search builder, select Common Data Element Resource and enter "umls" into the associated text box or add "umls[Common Data Element Resource]" to the search box at the top of the page and click on the Search button
  - The Studies tab of the search results displays all of the studies that contain tagged variables
  - The Variables tab of the search results displays all of the dbGaP variables that are tagged with at least one UMLS term
  - Click on a variable name to see more information on the variable page

- In the search builder, select Common Data Element Term and enter the CUI of a UMLS term into the associated text box or add "C0005890[Common Data Element Term]" to the search box at the top of the page and click on the Search button
    - The Studies tab of the search results display all of the studies that contain variables tagged with that UMLS term
    - The Variables tab of the search results displays all of the dbGaP variables that are tagged with that UMLS term
    - Click on a variable name to see more information on the variable page

Back to top [8]

Faceted search

- In your web browser, visit the dbGaP faceted search page [11]
    - You can navigate here by visiting the Advanced Search link under Access dbGaP Data on the dbGaP home page
- Click on the Variables tab
- Under the Common Data Elements filter, check UMLS
    - This will display all of the dbGaP study variables that are tagged with a UMLS term
- For a given variable listed on the right, you can click on the UMLS link to go directly to the variable's information page with the full UMLS term name
- To search for variables tagged with a specific UMLS term, search for the term's CUI in the search box in the upper left corner of the page

Back to top [8]

Variable page

- You can access a Variable page for any dbGaP variable by clicking on the variable name in a search results list
- If the variable has been tagged with a UMLS term, the variable page will have a Terms Linked to this Variable section, which will contain a table displaying information for any UMLS terms linked to the given variable
    - You can follow a link from this table to a list of all variables tagged with the same UMLS term (an Entrez search results page)

Back to top [8]

The 65 phenotype tags mapped to UMLS concepts

The DCC has mapped each tag to the UMLS [2] CUI that most closely matches the tag definition.

Phenotype Domain	Tag name (phenotype concept)	Description	UMLS CUI	UMLS term
Demographics	Gender	Self-reported sex or gender identity	C0017249	Gender identity
Demographics	Race/ancestry/ethnicity	Self-reported race, ancestry or ethnicity	C1830369	Race or ethnicity
Anthropometry	Height	Standing body height measurement	C0005890	Body height

Phenotype Domain	Tag name (phenotype concept)	Description	UMLS CUI	UMLS term
Anthropometry	Weight	Body weight measurement	C0005910	Body weight
Anthropometry	BMI	Body mass index (weight divided by the square of height)	C1305855	Body mass index
Anthropometry	Waist circumference	Waist circumference measurement	C0455829	Waist circumference
Anthropometry	Hip circumference	Hip circumference measurement	C0562350	Hip circumference
Anthropometry	Waist-hip ratio	Ratio of waist to hip circumference	C0205682	Waist-hip ratio
Cardiovascular disease	CAC	Quantitative measure of coronary artery calcification (CAC) by computed tomography (CT) scan	C1611184	Calcification of coronary artery
Cardiovascular disease	Carotid IMT	Quantitative measure of carotid intima media thickness (IMT) by ultrasound	C1960466	Carotid Intima-Media Thickness
Cardiovascular disease	Myocardial infarction	Qualitative indicator of myocardial infarction (MI) status	C2926063	Myocardial infarction:Finding:Point in time:^Patient:Ordinal
Cardiovascular disease	Coronary angioplasty	Qualitative indicator of coronary angioplasty status	C3542439	presence of coronary angioplasty implant and graft (physical finding)
Cardiovascular disease	Coronary artery bypass graft	Qualitative indicator of coronary artery bypass graft (CABG) status	C0585628	Presence of coronary artery bypass graft
Cardiovascular disease	Heart failure	Qualitative indicator of heart failure status	C0018801	Heart failure
Blood pressure	Resting arm systolic BP	Quantitative measure of resting systolic blood pressure	C2039694	Systolic blood pressure at rest
Blood pressure	Resting arm diastolic BP	Quantitative measure of resting diastolic blood pressure	C2183311	Diastolic blood pressure at rest
Blood pressure	Hypertension	Qualitative indicator of hypertension status	C3843080	Hypertension or high blood pressure
Diabetes	Blood glucose	Quantitative measure of the concentration of glucose in blood.	C1320980	Blood glucose status
Diabetes	Insulin in blood	Quantitative measure of the concentration of insulin in blood.	C0428405	Insulin level result
Diabetes	HbA1c	Quantitative measure of the concentration of glycated hemoglobin (hemoglobin A1c, or HbA1c) in blood	C1261236	Hemoglobin A1c level result
Diabetes	Diabetes	Qualitative indicator of diabetes mellitus status	C0011849	Diabetes Mellitus
EKG/Arrhythmia	Atrial fibrillation/flutter	Qualitative indicator of atrial fibrillation or atrial flutter status	C0155709	Atrial fibrillation and flutter
EKG/Arrhythmia	QRS duration from EKG	Quantitative measure of QRS duration from electrocardiogram (EKG) data	C0429025	QRS complex duration (observable entity)
EKG/Arrhythmia	QT interval from EKG	Quantitative measure of QT interval from electrocardiogram (EKG) data	C0429028	QT interval feature (observable entity)

Phenotype Domain	Tag name (phenotype concept)	Description	UMLS CUI	UMLS term
EKG/Arrhythmia	PR interval from EKG	Quantitative measure of PR interval from electrocardiogram (EKG) data	C0520877	PR interval feature
EKG/Arrhythmia	Resting heart rate from EKG	Quantitative measure of heart rate (HR) from electrocardiogram (EKG) data	C2041121	electrocardiogram rhythm strip 3-lead heart rate (bpm)
EKG/Arrhythmia	LVH from EKG	Quantitative index of left ventricular hypertrophy (LVT) calculated from electrocardiogram (EKG) data	C0232306	Electrocardiogram: left ventricle hypertrophy (finding)
EKG/Arrhythmia	Pacemaker	Qualitative indicator of pacemaker implant status	C1533090	Pacemaker observable
Hematology & Hemostasis	Hematocrit	Quantitative measure of the fraction of blood volume that is composed of red blood cells	C0518014	Hematocrit level
Hematology & Hemostasis	Hemoglobin	Quantitative measure of hemoglobin concentration in blood	C0019029	Hemoglobin concentration result
Hematology & Hemostasis	Platelet count	Quantitative measure of platelet cell number per volume of blood	C1287267	Finding of platelet count
Hematology & Hemostasis	Red blood cell count	Quantitative measure of red blood cell number per volume of blood	C1287262	Finding of red blood cell count
Hematology & Hemostasis	White blood cell count	Quantitative measure of white blood cell number per volume of blood	C0427512	White blood cell count laboratory result
Hematology & Hemostasis	Fibrinogen in blood	Quantitative measure of fibrinogen concentration in blood	C1318051	Fibrinogen level result
Hematology & Hemostasis	Factor VII	Quantitative measure of factor VII activity or concentration in blood	C0427593	Factor VII level
Hematology & Hemostasis	Factor VIII	Quantitative measure of factor VIII activity or concentration in blood	C0457997	Factor VIII level
Hematology & Hemostasis	von Willebrand factor	Quantitative measure of von Willebrand factor (vWf) activity or concentration in blood	C0427585	vWF - von Willebrand factor level result
Hematology & Hemostasis	VTE	Qualitative indicator of venous thromboembolism (VTE) status	C1861172	venous thromboembolism
Inflammation	CRP in blood	Quantitative measure of C-reactive protein (CRP) concentration in blood	C0428528	C-reactive protein level
Inflammation	Interleukin 6 in blood	Quantitative measure of Interleukin 6 (IL-6) concentration in blood	C0366888	Interleukin 6:ACnc:Pt:Ser/Plas:Qn
Kidney function	Creatinine in blood	Quantitative measure of creatinine concentration in blood	C0428279	Finding of creatinine level
Kidney function	Cystatin C in blood	Quantitative measure of cystatin C concentration in blood	C1445989	Blood cystatin C measurement
Kidney function	Albumin-creatinine ratio in urine	Quantitative measure of the ratio of albumin concentration to creatinine concentration in urine	C1318330	Urine albumin:creatinine ratio result

Phenotype Domain	Tag name (phenotype concept)	Description	UMLS CUI	UMLS term
Kidney function	GFR	Quantitative measure of glomerular filtration rate (GFR)	C4524116	Estimated GFR
Lipids	LDL in blood	Quantitative measure of low-density lipoprotein (LDL) cholesterol concentration in blood	C2603388	Cholesterol.in LDL [Mass or Molecules/volume] in Serum or Plasma
Lipids	HDL in blood	Quantitative measure of high-density lipoprotein (HDL) cholesterol concentration in blood	C2603387	Cholesterol.in HDL [Mass or Molecules/volume] in Serum or Plasma
Lipids	Triglycerides in blood	Quantitative measure of triglyceride concentration in blood	C0428475	Triglyceride level - finding
Lipids	Total cholesterol in blood	Quantitative measure of total cholesterol concentration in blood	C0201950	Cholesterol measurement test
Lung	FVC	Quantitative measure of forced vital capacity (FVC) measured using spirometry	C3714541	Forced Vital Capacity
Lung	FEV1	Forced expiratory volume in 1 second (FEV1) measured using spirometry	C0802965	Volume^at 1.0 s post forced expiration:Volume:Point in time:Respiratory system:Quantitative
Lung	Asthma	Qualitative indicator of asthma status	C1272273	Asthma finding
Lung	Asthma severity	Qualitative and quantitative measures of asthma severity	C0581122	Asthma severity
Lung	COPD	Qualitative indicator of Chronic Obstructive Pulmonary Disease (COPD) status	C0024117	Chronic Obstructive Airway Disease
Sleep	Sleep apnea	Qualitative indicator of sleep apnea status	C0037315	Sleep Apnea Syndromes
Sleep	AHI	Quantitative measure of Apnea-Hypopnea Index (AHI), a measure of sleep apnea severity	C2111846	Apnea-hypopnea index procedure
Smoking	Cigarette smoking	Qualitative and quantitative measures of cigarette smoking history and cigarette smoking status	C1519384	Smoking History
Stroke	Ischemic stroke	Qualitative indicator of ischemic stroke status	C0948008	Ischemic stroke
Stroke	Hemorrhagic stroke	Qualitative indicator of hemorrhagic stroke status	C0553692	Brain hemorrhage
Stroke	Other stroke	Qualitative indicator of other/unspecified stroke status	C4489300	Stroke- other stroke
Supporting phenotypes	Age at enrollment/collection	Participant age at enrollment or age at which data or biosamples were collected	C0001779	Age
Supporting phenotypes	Subcohort	Categorical indicator of membership in different study subcohorts (i.e. participant subsets within a study who were recruited at different times and/or enrolled in different sub-studies)	C0599755	Cohort

Phenotype Domain	Tag name (phenotype concept)	Description	UMLS CUI	UMLS term
Supporting phenotypes	Clinic visit	Categorical indicator of the clinical visit (exam, year, etc.) during which observations, measurements and/or other data/biosample collections were made	C0008952	Clinic Visits
Supporting phenotypes	Fasting	Qualitative or quantitative indicator describing or quantifying fasting status prior to blood draw for any blood sample-derived measurement	C1976106	Fasting Status
Supporting phenotypes	Geographic site	Categorical indicator of the clinic, recruitment site, and/or field center at which participants were recruited, and/or where data or biosamples were collected.	C2828208	Locality
Supporting phenotypes	Medication/supplement use	Qualitative or quantitative indicator of use of any kind of medication, vitamin, mineral, or dietary supplement	C0240320	medication use

CSV [12]

Back to top [8]